

伝統芸能実演家の動的データベースの作成

坂部 裕美子

公益財団法人 統計情報研究開発センター 研究開発本部

Creation of the Dynamic Database about the Traditional-Performing-Arts Performers

Yumiko Sakabe

Research-and-Development Headquarters,
Statistical Information Institute for Consulting and
Analysis

要旨：

伝統芸能の実演者に関する長期データ中から、ある任意の年次時点で活動していた実演家のみを抜き出し、比較用の属性を付加して動的データベースを作成するプログラムを開発した。その適用例を紹介する。

キーワード：arrayステートメント、mergeステートメント、updateステートメント、rankプロシジャ

1 プログラムの必要性

- 筆者は、伝統芸能の上演傾向等を年度ごとに集計・比較するという研究をしている
- 年度ごとの公演概況集計結果の参照用資料として、当時の実演家の付帯情報をまとめたDBが必要
- 実演者リストは、現実的には単年度版が市販されることが多く、観客としてはその方が使いやすい。しかし、データ整備・更新作業面からは、積年版の単一ファイルの方が好都合

作業イメージ

VIEWTABLE: Work.All_data

	player_ID	born	enter	retire	name	class
1	1	1910	1930	1985	ぜん二	3
2	2	1920	1940	1990	ぜん三	3
3	3	1930	1950	.	ぜん四	3
4	4	1940	1960	.	ぜん五	2
5	5	1950	1970	2000	ぜん六	2
6	6	1960	1980	.	ぜん七	1
7	7	1970	1990	.	ぜん八	3
8	8	1980	1995	.	ぜん九	3
9	9	1985	2000	.	ぜん十	3
10	10	1990	2005	.	まえ一	3
11	11	1910	1935	1990	まえ二	3
12	12	1920	1945	2000	まえ三	3
13	13	1930	1955	1995	まえ四	2
14	14	1940	1955	.	まえ五	2
15	15	1950	1970	.	まえ六	1
16	16	1960	1980	.	まえ七	1
17	17	1970	1995	.	まえ八	1
18	18	1980	2000	.	まえ九	1
19	19	1985	2005	.	まえ十	1
20	20	1990	2010	.	さき一	1
21	21	1910	1930	1980	さき二	1
22	22	1920	1935	.	さき三	1
23	23	1930	1955	.	さき四	1
24	24	1940	1960	2005	さき五	1
25	25	1950	1980	.	さき六	1
26	26	1960	1990	.	さき七	1
27	27	1970	1985	.	さき八	1
28	28	1980	2000	.	さき九	1
29	29	1985	2005	.	さき十	1
30	30	1990	2005	.		1

VIEWTABLE: Work.Db_1990

	player_ID	name	age	class
1	2	真二	70	3
2	3	真三	60	3
3	4	真四	50	3
4	5	ふた五郎	40	2
5	6	ふた六	30	2
6	7	ぜん七	20	1
7	11	打一	80	3
8	12	打二	70	3
9	13	打三	60	3
10	14	打四	50	3
11	15	つめ五	40	2
12	16	まえ六	30	1
13	22	師二	70	3
14	23	師三	60	3
15	24	はい四	50	2
16	25	はい五	40	2
17	26	さき六	30	1
18	27	さき七	20	1

2005

2005
2010
1930	はい一	1955	師一	1970
1935	はい二	1965	師二	1975
1955	はい三	1975	師三	1985
1960	はい四	1980	師四	1995
1980	はい五	1985	.	.
1990
1985
2000
2005
2005

(注)この図のPlayer_IDは作業用の仮番号

2 共通作業の存在

- 実際に「歌舞伎」「落語」「宝塚」の実演家DBを開発していく過程で、すべてに共通する作業過程があり、プログラムを共有できるのではないかと考えた
 - 襲名および改名...該当年度の名前で表示する
 - 昇進...該当年度の階級で表示する
- 「共通プログラム」と「分野ごとの独自指標の算出」を経て、長期DBから特定の年度の実演家DBを抽出表示する

3 共通プログラム

- ① 全演者データから該当年度に活動していた者のみを抽出し、年齢を計算する

```
%let year=1990;
data db_&year;
  set all_data;
  format player_ID name age class <F1><F2>;
  if retire ne . and &year > retire then delete;
  if enter > &year then delete;
  %determ
  age= &year-born;
  keep player_ID name age class <F1><F2>; run;
```

Player_ID...各演者のユニークID
落語＝入門順
宝塚＝成績順

② その当時の階級と名称で表示する

```
%macro determ;
  class=0;
  %do i=1 %to 5;
    if name_&i ne " and shumei_&i=<&year then
      name=name_&i;
    if shoshin_&i=<&year then
      class+1;
  %end;
%mend;
```


4 分野特化プログラム①—歌舞伎

- 共通プログラム中の要修正箇所
 - 階級昇進なし、改名(襲名)あり
- 独自項目(実演者DBからの付加項目)
 - 親は誰か(将来的には「歌舞伎俳優か」も考慮する)
 - 特定の顕彰制度についての受賞歴
- 独自算出指標(DB内で算出するもの)
 - 家の格(受賞歴を使用)

「家の格」を定義

- 「芸術院会員」「人間国宝」「文化功労者」「文化勲章」のうち2つ以上の顕彰歴のある俳優は「名家」と考えたい(=ランク1)
- 上記の条件を満たす俳優の息子は「特別視」したい(=ランク2)
- これらの条件に該当する俳優にフラグを付ける
 - この「家の格」を使用した集計を、過去のSASユーザー総会で報告済み

(1) 4賞受賞者を抽出

```
data ds01;  
  set db_&year (keep=player_ID prize1-prize4);  
  array prz{4} prize1-prize4;  
  flg01=0;  
  do i=1 to 4;  
    flg01+(prz{i} ne . and prz{i} =< &year );  
  end;  
  if flg01>=2 then output;  
run;
```

「該当年次における受賞状況」とすることで、
役者自身の「ランク2」から「ランク1」への
昇格を反映できる
(例:菊五郎7)

(2) 該当者へフラグ付け

1. 親が該当するケースへ mergeする

```
data ds01;  
  set ds01(keep=player_ID);  
  prz01=2; run;  
proc sort data=db_&year;  
  by shisho;  
data ds02_1;  
  merge db_&year ds01  
  (rename=(player_ID=shisho));  
  by shisho;  
run;
```

2. 当人が該当するケースへ updateする

```
data ds01;  
  set ds01;  
  prz01=1;run;  
proc sort data=db_&year;  
  by player_ID;  
data ds02_2;  
  update ds02_1 ds01;  
  by player_ID;  
run;
```

両方に該当するケースは、
より上位の1に上書きされる

5 分野特化プログラム②ー落語

- 共通プログラム中の要修正箇所
 - なし(階級、名称、年齢すべて処理が必要)
- 独自項目
 - 師匠は誰か
- 独自算出指標
 - 入門以降の経過年数
 - 真打に抜擢昇進したか
 - 弟子／兄弟弟子の人数

抜擢昇進フラグ付け

- 落語家の階級 = 前座 / 二つ目 / 真打
 - 抜擢昇進: **二ツ目**の落語家が、先輩を飛び越えて**真打**に昇進するケース
 - 将来性を高く見込まれて抜擢が行われることが多い。実際に、昇進後の寄席の出番は、そうでないものより有意に多い
 - 「抜いた人数の多寡」よりも「抜擢されたか否か」の方がその後の活躍への影響が大きい
- 過去のSASユーザー会で分析報告済み

1. 未昇進者の真打昇進年に仮の値を入れる

```
data ds01;  
  set all_data(keep=player_ID shumei_3);  
  if shumei_3=. then shumei_3=2999;  
run;
```

入門後当該年までの死亡・廃業者も考慮し、正確な抜擢状況を把握する
(今後の研究で使用する可能性あり)

2. 真打昇進順位が入門順位より上の人にフラグを付ける

```
proc rank data=ds01 out=ds02 ties=low;  
  var shumei_3; ranks shoshin;  
run;  
data ds02; set ds02; if shoshin<player_ID then flg01=1;  
run;
```

弟子・兄弟弟子について

- 落語家のライフサイクルを考察する上で有意義な指標たりうると考えられる
- 弟子がいる ⇨ (1) ある程度成功している
(2) 後進の育成期に入っている
- 「弟子の数」や「自身が惣領弟子であるか否か」も指標の一つとなり得る？（今後の検証が必要）

1. 兄弟弟子の連番をふる

```
proc rank data=db_&year out=ds03_1 ties=low ;  
var enter; ranks deshi_no ; by shisho;  
run;
```

2. 自分の弟子の人数を付与する

```
proc freq data=db_&year; tables shisho / noprint out= ds04 ; run;  
data ds03_2; merge ds03_1 ds04  
(drop=percent rename=(shisho=player_ID count=deshi));  
by player_ID; run;
```

6 分野特化プログラム③ー宝塚

- 共通プログラム中の要修正箇所
 - 階級昇進なし、年齢なし、改名あり(僅かに存在)
- 独自項目
 - 所属する組、男役/娘役の別
 - トップ就任の有無および就任年
 - 新人公演の主演および本公演以外の主演状況
- 独自算出指標
 - 入団以降の経過年数
 - 本公演以外の公演種別主演回数

組替え情報の集計

- 過去の組替え回数は、その後の活動内容と関連があると考えられるので、別途項目として立てておく

この部分の作業がこれ以降で繰り返しになるので、マクロを定義する

```
data ds01;
set db_&year (keep=player_ID c_class_year: );
array c_class{*} c_class_year: ;
c_class_t=0;
do i=1 to dim(c_class);
  c_class_t+(c_class{i} ne . and c_class{i} =< &year );
end;
run;
```

公演主演状況

- 新人公演・バウホール公演・その他の劇場の主演経験歴を集計する

```
data ds02;  
set db_&year (keep=player_ID shin_top: vow_top: other_top: );
```

```
%gsum(shin)  
%gsum(vow)  
%gsum(other)
```

```
run;
```

```
%macro gsum(var);  
array &var{*} &var_top_: ;  
&var_main=0;  
do i=1 to dim(&var);  
  &var_main+(&var{i} ne . and &var{i} =< &year );  
end;  
%mend;
```

まとめ

- 収録データの規模は、「全データ」で最大4500件程度、「抽出データ」で最大400件程度(うち、独自指標の算出が必要なのは数十件)の見込みなので、処理時間の問題なく実用化できると考えられる
- Outputに関しては未開拓だが、もともと結果閲覧が目的なので、将来的に着手できればと考えている
- この分野の研究にSASを使っている人は皆無だと思うので、珍奇な事例報告としてご紹介しました。まだプログラムに改良の余地はあると思います。